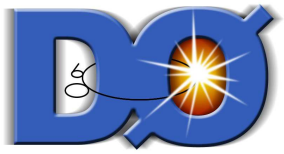
Karlsruhe Regional Analysis Center (GridKA)



Daniel Wicke
(Bergische Universität Wuppertal)

Outline

- Introduction
- GridKa
- RAC Prototype at GridKa
- Analyses run at GridKa
- Outlook and Summary

Introduction

Relieve central computing system by worldwide distribution (DØNote 3984)

Regional Analysis Centers (RACs)

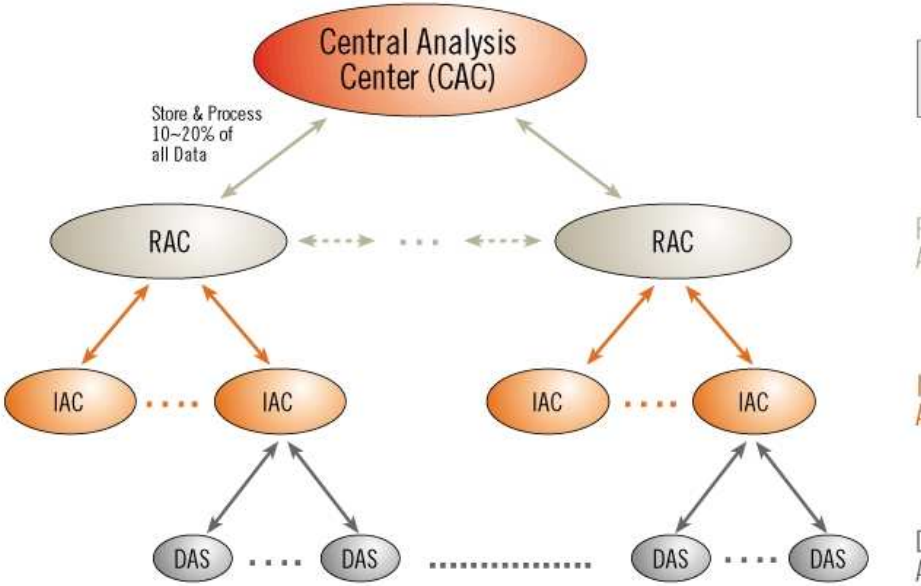
- Allow full physics analysis.

  ⇒ Hold all Thumbnails.
  ⇒ Provide computing power to process these.

- Allow distributed reprocessing.

  ⇒ Hold 10% of full DST.

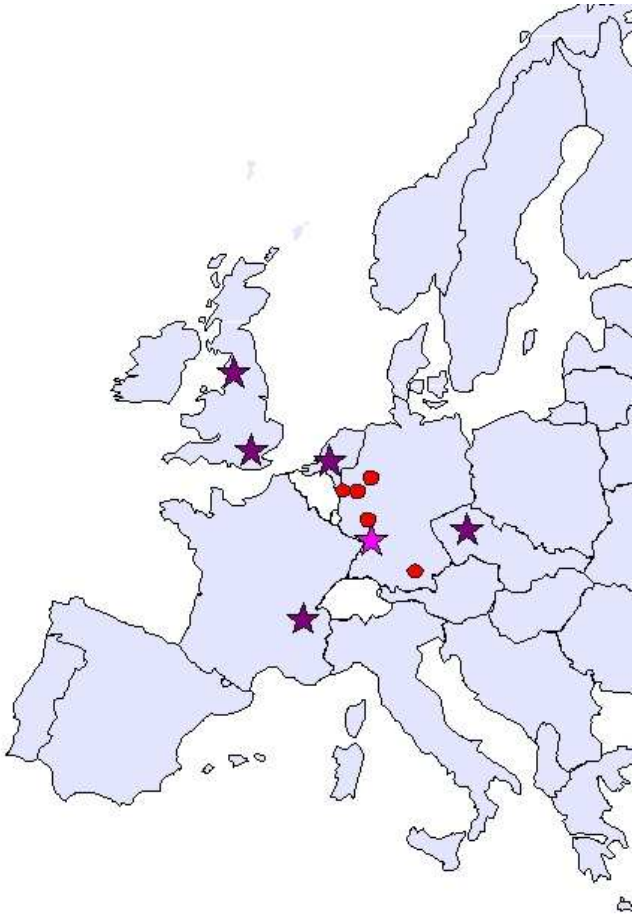
- Serve institutes.



- As such a major building block for the DØ Grid.

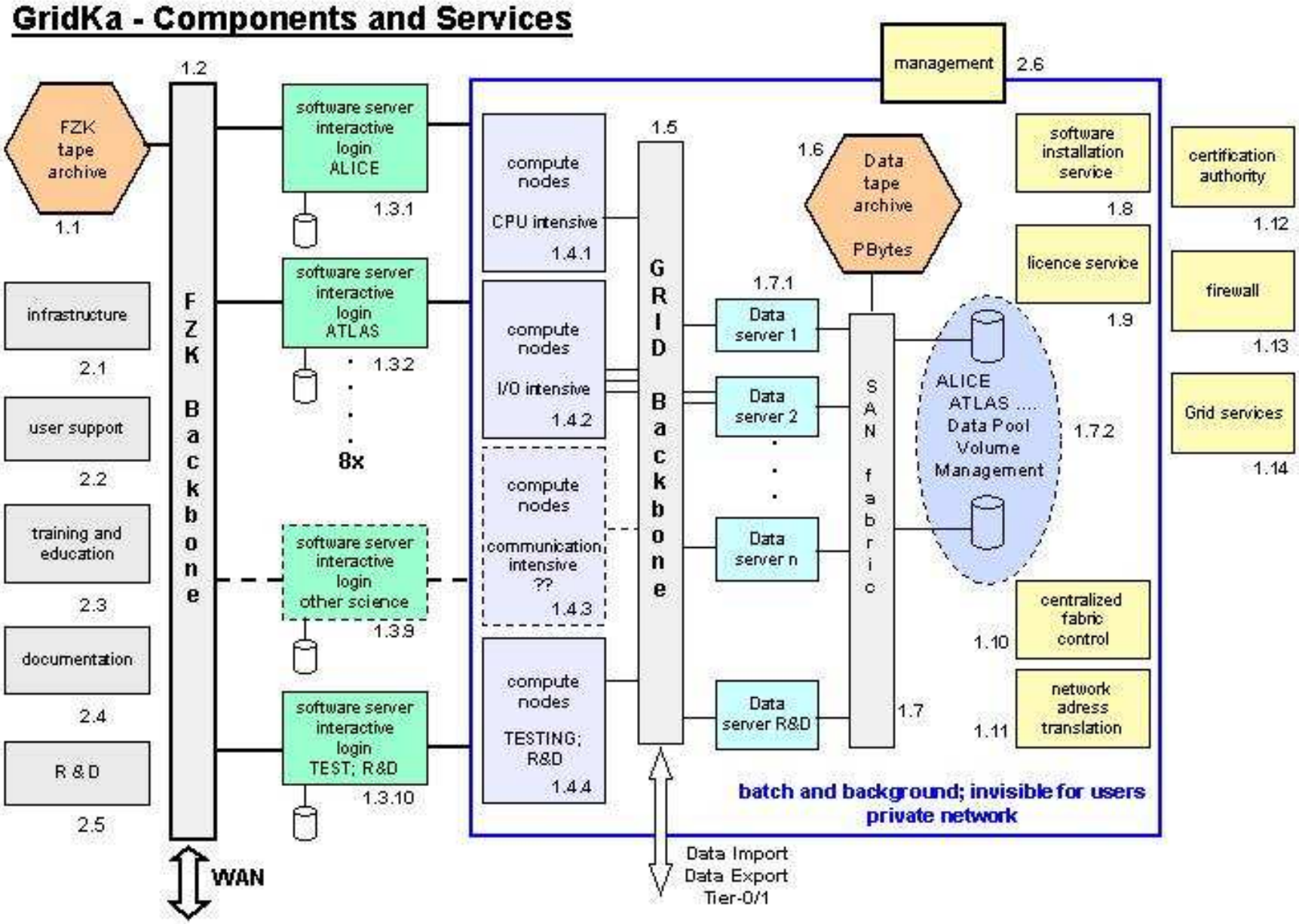
Grid Computing Centre Karlsruhe: GridKa

- located at Forschungszentrum Karlsruhe (FZK).
- established in 2002.
- centre for Grid development.
- regional data and computing centre.
- 8 HEP experiments:
  Alice, Atlas, Babar, CDF, CMS,
  Compass, DØ and LHCb.



Political Structure

- Overview Board:
  FNAL represented by Peter Mättig, Wuppertal.
- Technical Advisory Board:
  DØ represented by Christian Zeitnitz, Mainz, and D.W., Wuppertal.


GridKa - Components and Services
management
2.6
1.2
FZK tape archive
1.1
software server interactive login ALICE
1.3.1
software server interactive login ATLAS
1.3.2
8x
software server interactive login other science
1.3.9
software server interactive login TEST; R&D
1.3.10
FZK Backbone
infrastructure
2.1
user support
2.2
training and education
2.3
documentation
2.4
R & D
2.5
WAN
compute nodes CPU intensive
1.4.1
compute nodes I/O intensive
1.4.2
compute nodes communication intensive ??
1.4.3
compute nodes TESTING; R&D
1.4.4
1.5
GRID Backbone
1.6
Data tape archive PBytes
1.7.1
Data server 1
Data server 2
Data server n
Data server R&D
SAN fabric
1.7
ALICE ATLAS .... Data Pool Volume Management
1.7.2
software installation service
1.8
licence service
1.9
centralized fabric control
1.10
network adress translation
1.11
certification authority
1.12
firewall
1.13
Grid services
1.14
batch and background; invisible for users
private network
Data Import
Data Export
Tier-0/1

Hardware status

DØ Software Server (d0.fzk.de)

Dual Pentium III 1.26GHz, RAM: 2GB, HD: 4× 75GB IDE-Raid,
100Mbit/s Ethernet to FZK Backbone, Gbit/s Ethernet to Grid Backbone.

Compute Servers (as of February 2003, next update April 2003)

95× dual Pentium III 1.2GHz, 1GB RAM, HD 40GB IDE, 100Mbit/s Ethernet
68× dual Xeon 2.2GHz, 1GB RAM, HD 40GB IDE, 100Mbit/s Ethernet

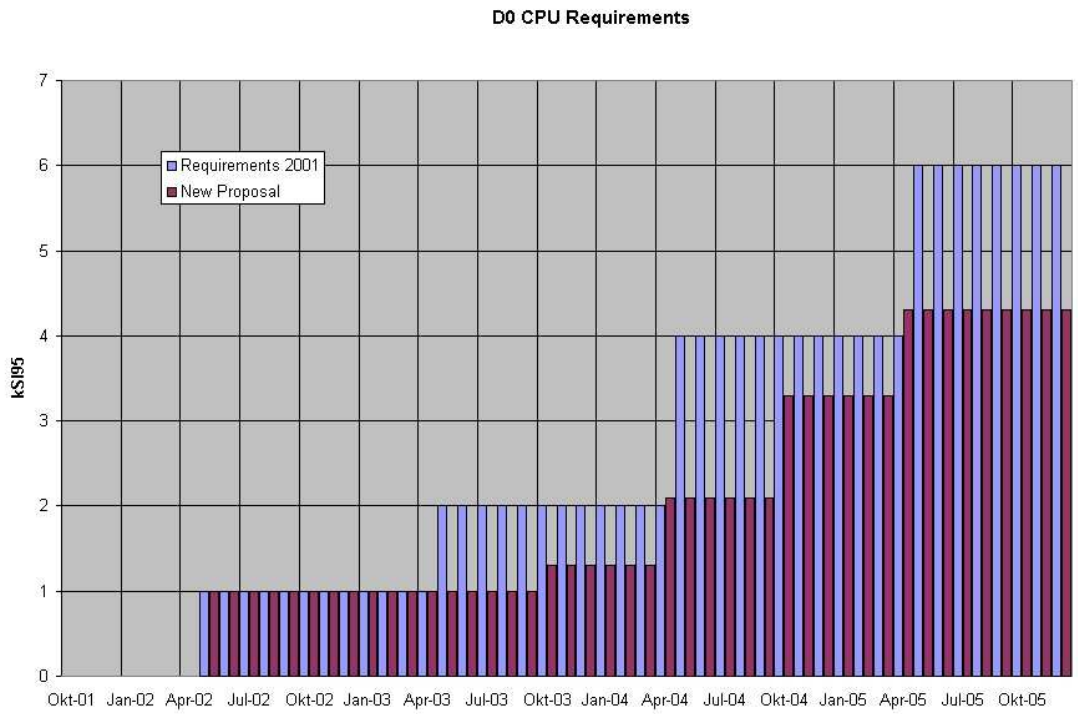
DØ has requested 6% of this CPU power.

Data Pool (as of February 2003, next update April 2003)

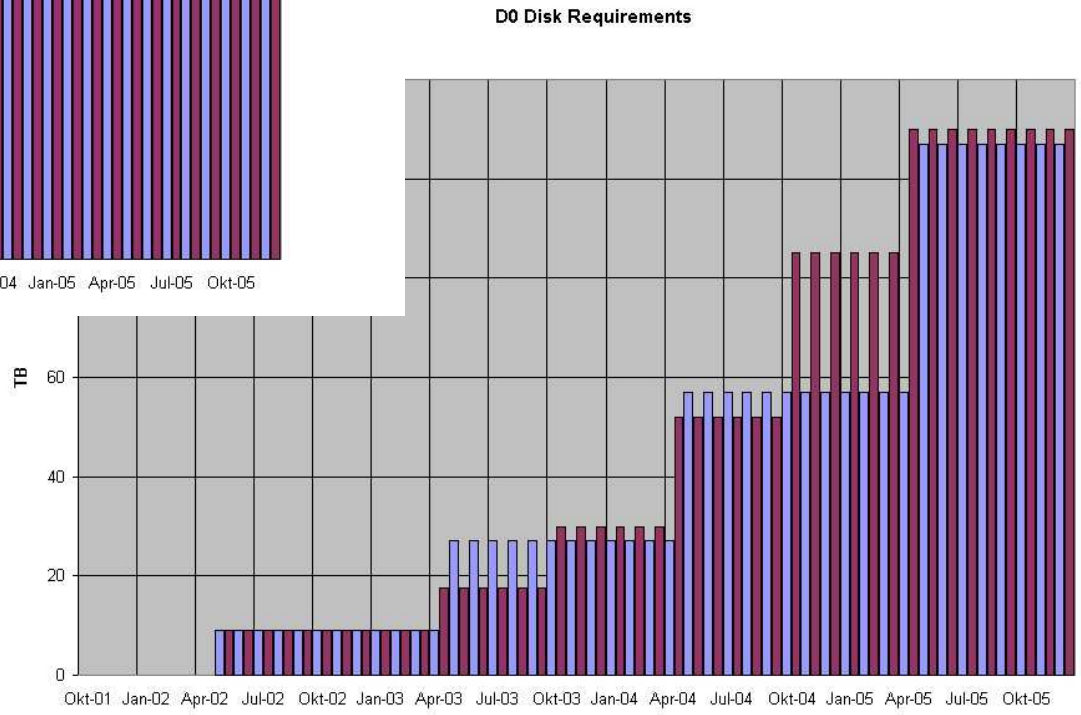
Several IDE Raid Arrays: 47TB brutto $\simeq$ 40TB netto.
Of these 3.4TB are available for DØ (0.4TB more than requested for 2002).

IBM Powderhorn Tape Robot with Tivoli Storage Manager: $\mathcal{O}$(100TB).

Prospected Developement (DØ only)



← CPU (SPECInts)



Diskspace →

Software status

Cluster Setup

- Compute nodes have private IP adresses.
- Compute nodes see all user and data disk areas via nfs.
- Batch system: OpenPBS v2.2p5-7

DØ Software Server (`d0.fzk.de`)

OS: Redhat 7.2 (Enigma), Kernel 2.4.18

Extensions

- ups/upd
- DØ-Software — p10, p11, p12 and p13 available.
- SAM Station — v4_2_1_31 (nfs-shared cache)

Root access is in the hand of GridKa.

RAC Prototype

Goals

- Proof of principle for RAC concept.
  Provide a working analysis environment for DØ-Germany.
- Check needed resources and its scalability (mostly person power and network).
- Check D0-software compatibility.

Specifications

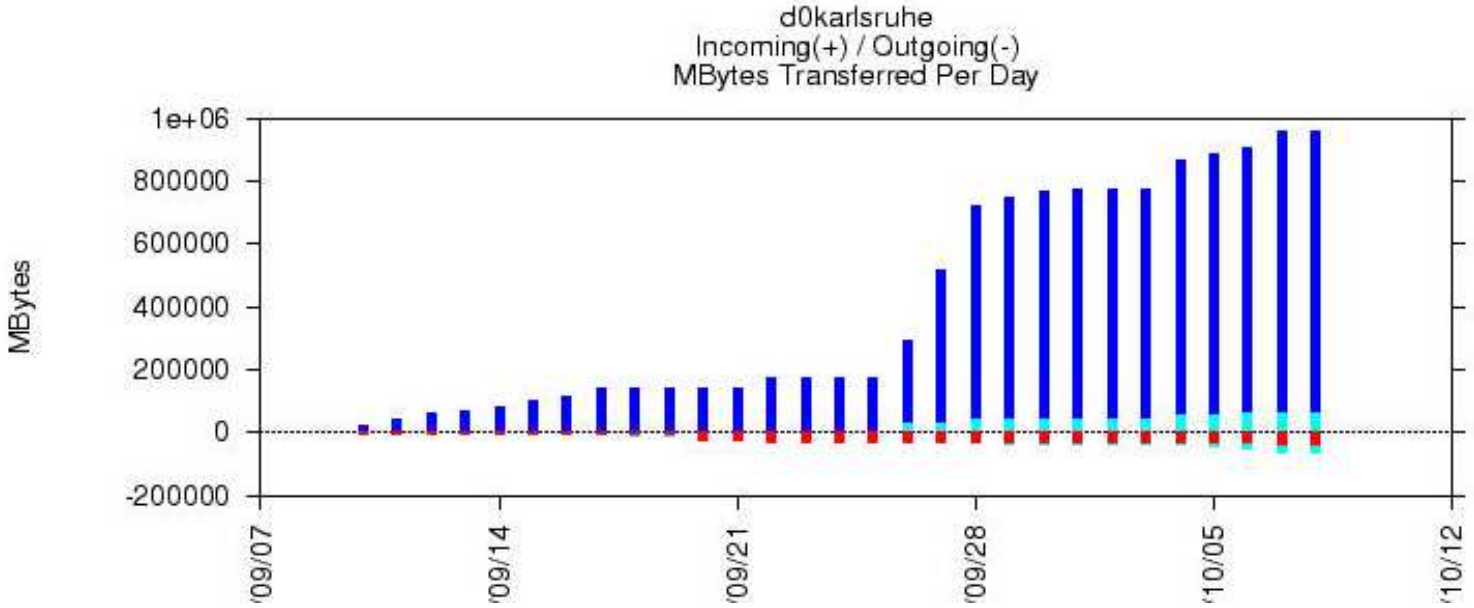
The prototype should implement the following major features of a full RAC:

- Continuously and immediate transport of thumbnails
  from Fermilab to GridKa disks.
- Fetching files available at the prototype from associated institutes.
- Automatic installation of DØ-software updates.

Status of Prototype at GridKa

1. Continuously and immediate transport of thumbnails

- Up and running since end of August 2002: a cronjob every 2h.
- $\simeq$2.5TB thumbnails transported (cache size 2.6TB).
- Transfer parameters (e.g. speed, retry rate, failure rate, etc.) are available on various web pages.

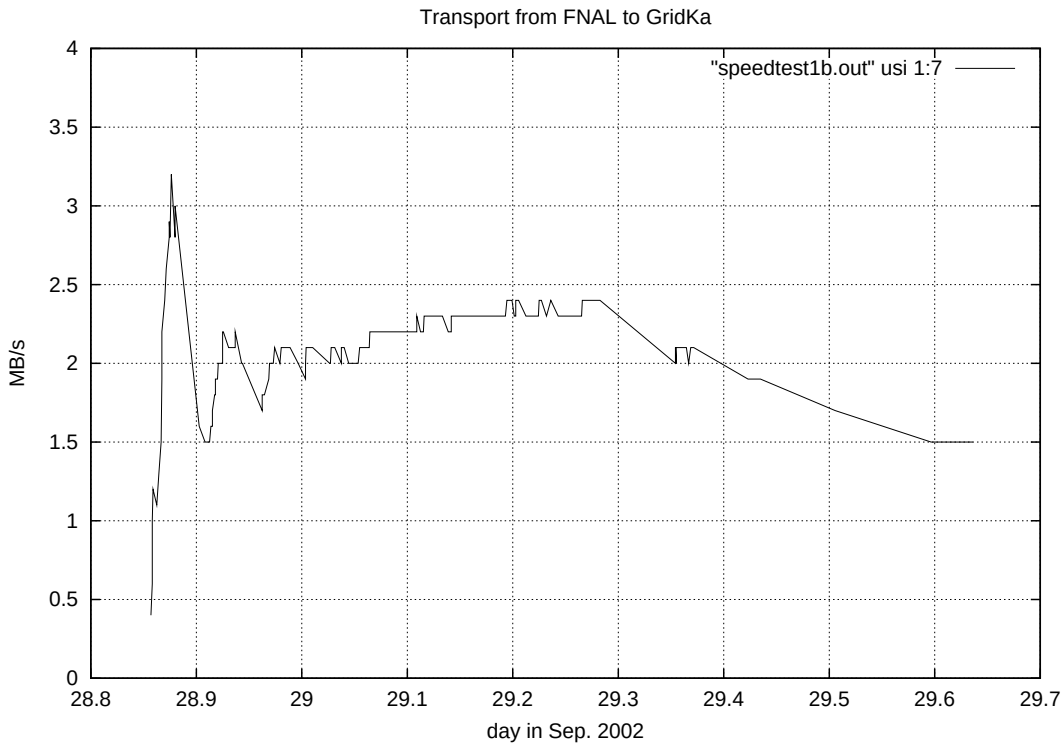


Results: Transfer Speed

Integrated size of arriving files over time:

2–3MB/s = 15–25MBit/s (averaged over 7 hours).

larger gaps in the transport decreased effective speed thereafter.



$\Longrightarrow$ (At that time) limited by FZK connection (32MBit/s)

Status of Prototype at GridKa (II)

2. Fetching files available at the prototype from associated institutes.

   - Implemented in SAM stations of Munich and Wuppertal:

     ```
     --routing-station=\.\*::d0karlsruhe
     --routing-user=wuppertal
     --routing-group=dzero
     ```

     (after opening the GridKa firewall for the corresponding sam-stations)

   - Files missing at GridKa are routed through GridKa to the requesting institute.

3. Automatic installation of DØ-software updates.

   - Started by Christian Autermann, Aachen.

Software Problems

- SAM doesn't allow to share cache between nodes. (A. Baranowski, FNAL)
- SAM suddenly stopped to operate on 23rd Sep. at 2:45am. (C. Schmitt, Wtal)
- Transport Problems (Christian Schmitt, Wuppertal)
- `mc_runjob` (David Meder, Mainz)
- DØTools (Thomas Nunnemann, Munich)
- Luminosity (open issue)
- Software access through `d0cvs` (Marc Hohlfeld, Mainz)
- Tape access for SAM (D.W., Wuppertal)
- CAB induced transport errors in SAM (Christian Schmitt, Wuppertal)

Thanks to the GridKa and SAM-admin teams for the support.

Example: CAB induced transport problems

The launch of CAB induced unexpected problems in Karlsruhe:

Symptom

- Large number of transfer errors/undelivered files.

Causes

- GridKa sam-station tried to get files directly from CAB.
- This failed because the CAB-nodes aren't known to our firewall.

Solutions

- Opening the firewall isn't feasable.
- Route files from CAB through central-router (`d0mino`).

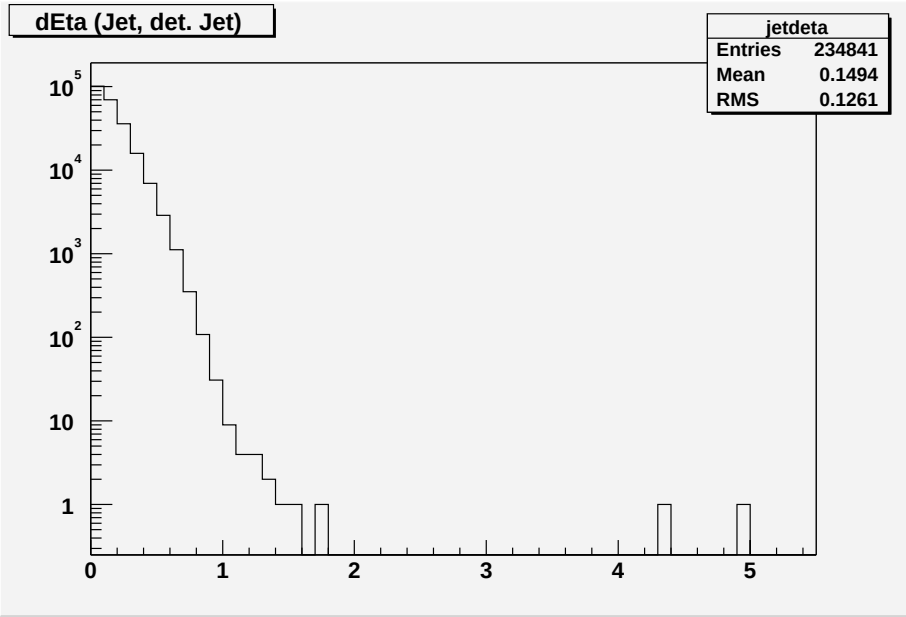
Example of how changes at Fermilab can affect remote stations.

$\Longrightarrow$ We should aim to remove/avoid these kind of dependencies:

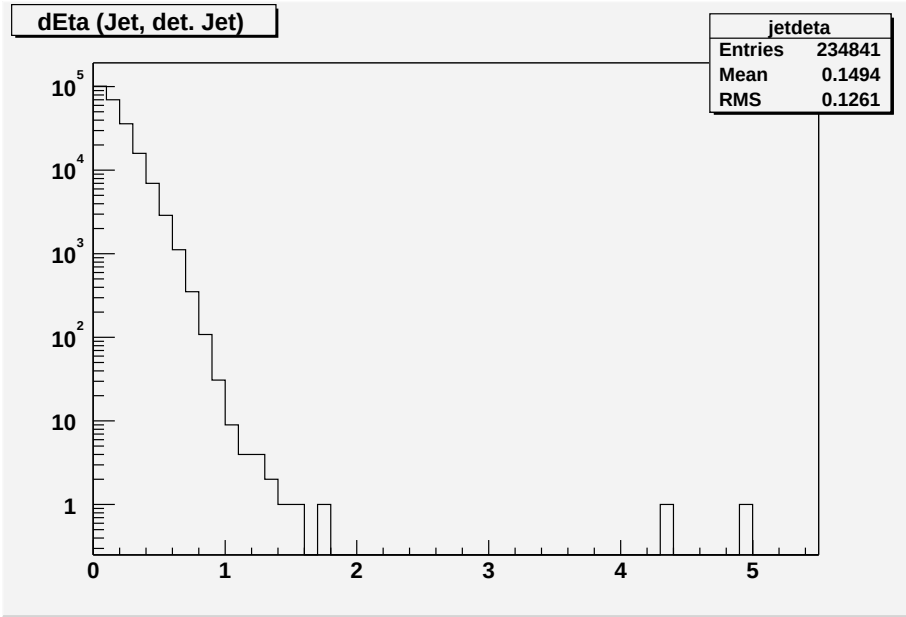
Think globally, avoid site specific paths, code and libraries.

Analyses run at GridKa

Comparison of results produced at different sites



`Clued0`



`GridKa`

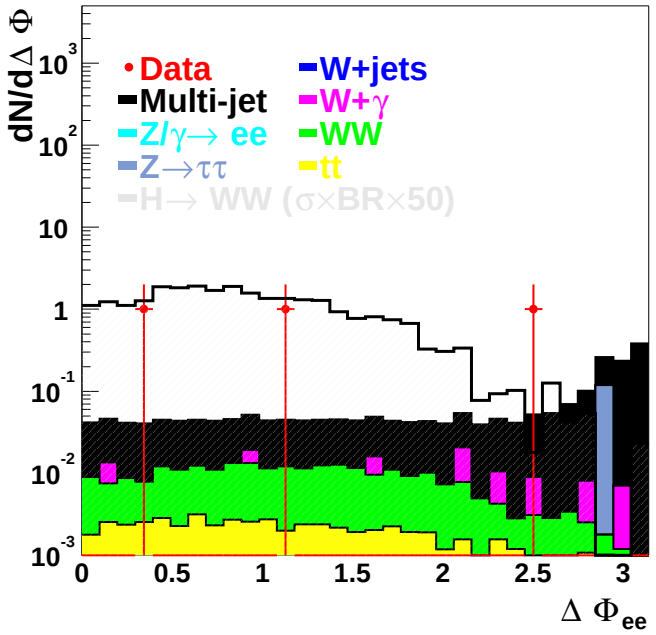
Produced using `top_analyze v00-02-06` on `Clued0` and on `GridKa` on a small set of Thumbnails.

Similar crosschecks by Johannes Elmsheuser, Munich, include Munich.

Physics results

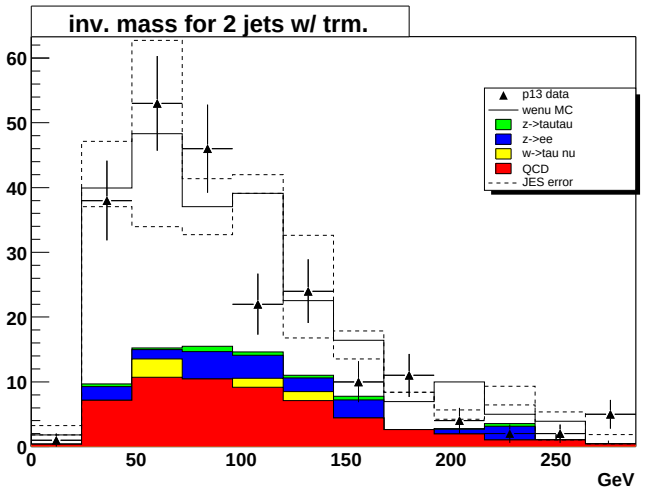
- $H \longrightarrow W^+W^-$
  (Johannes Elmsheuser, Munich,
  and Marc Hohlfeld, Mainz)
- W+jets
  (David Meder, Mainz)

Both analyses produce results for Moriond.



MC Productions

- 1.2 million MC events for the top group
  (Markus Klute, Bonn)
- Vecbos vs. Alpgen comparisons
  (Christian Schmitt, Wuppertal)
- ...



Summary

- GridKa (Grid Computing Center Karlruhe) is the German Grid Center.
  - DØ is represented in the technical and political boards.
- Prototype DØ Regional Analysis Center has been successfully set up.
  - Continuous transport of TMB.
  - Routing/intermediate caching for institutes.
  - Users.
- Starts to unburden FNAL computing resources.

The RAC concept has proven its value.

Outlook

- Evolve to a fully functional *analysis* center ("production mode").
- Go for DØGrid.
- Check out data reprocessing abilities.